

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



538204

(43) International Publication Date
24 June 2004 (24.06.2004)

PCT

(10) International Publication Number
WO 2004/053795 A2

(51) International Patent Classification⁷: **G06T 7/00**

(21) International Application Number:
PCT/IB2003/005773

(22) International Filing Date: 8 December 2003 (08.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/432,539 11 December 2002 (11.12.2002) US

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LEE, Mi-Suen** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **LIN, Yun-Ting** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **TRAJKOVIC, Miroslav** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **PHILOMIN, Vasanth** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(74) Common Representative: **KONINKLIJKE PHILIPS ELECTRONICS N.V.**; Intellectual Property & Standards, c/o THORNE, Gregory, L., P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

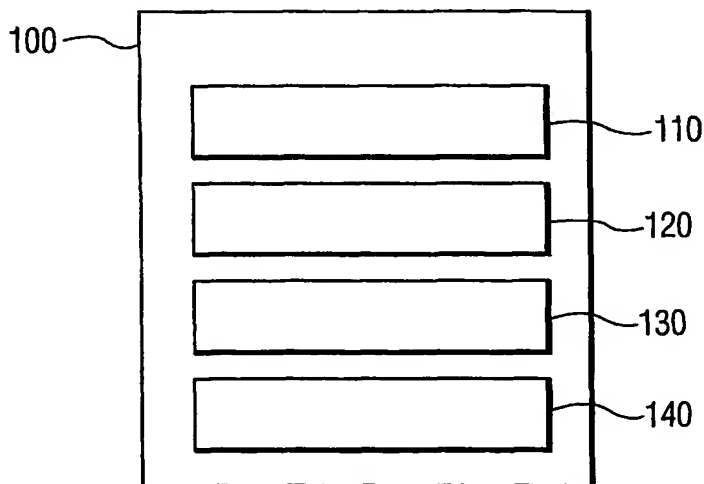
— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for all designations

Published:

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR CORRECTING A HEAD POSE IN A VIDEO PHONE IMAGE



(57) Abstract: An image processing system (250) and method (300) are disclosed for correcting a head pose in a video phone image, so that a frontal view is presented on a display. A disclosed head pose corrector (250) estimates the orientation of a head pose and adjusts the orientation of the head pose, if necessary, to present a frontal view. The orientation of the head pose is adjusted by generating a three dimensional model of the face surface and adjusting the orientation of the three dimensional face model to provide the desired frontal view. The head pose corrector (250) may be included in a video phone (100) to correct the head pose of transmitted or received images (or both) or may be included in a server on a network to automatically adjust the head shots of one or more participants to a video phone communication.

10/528204
Rec'd PCT/PTO 09 JUN 2005

WO 2004/053795 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
24 June 2004 (24.06.2004)

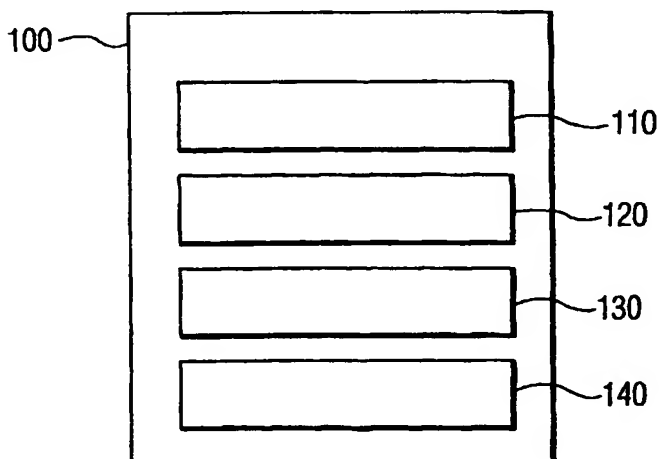
PCT

(10) International Publication Number
WO 2004/053795 A3

- (51) International Patent Classification⁷: **G06T 7/00**, 15/20, H04N 7/14
- (21) International Application Number: **PCT/IB2003/005773**
- (22) International Filing Date: 8 December 2003 (08.12.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/432,539 11 December 2002 (11.12.2002) US
- (71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]**; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **LEE, Mi-Suen** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **LIN, Yun-Ting** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **TRAJKOVIC, Miroslav** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US). **PHILOMIN, Vasanth** [US/US]; P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).
- (74) Common Representative: **KONINKLIJKE PHILIPS ELECTRONICS N.V.**; Intellectual Property & Standards, c/o THORNE, Gregory, L., P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Declaration under Rule 4.17:**
— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for all designations
- Published:**
— with international search report
- (88) Date of publication of the international search report:
16 June 2005

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR CORRECTING A HEAD POSE IN A VIDEO PHONE IMAGE



(57) Abstract: An image processing system (250) and method (300) are disclosed for correcting a head pose in a video phone image, so that a frontal view is presented on a display. A disclosed head pose corrector (250) estimates the orientation of a head pose and adjusts the orientation of the head pose, if necessary, to present a frontal view. The orientation of the head pose is adjusted by generating a three dimensional model of the face surface and adjusting the orientation of the three dimensional face model to provide the desired frontal view. The head pose corrector (250) may be included in a video phone (100) to correct the head pose of transmitted or received images (or both) or may be included in a server on a network to automatically adjust the head shots of one or more participants to a video phone communication.

WO 2004/053795 A3



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

National Application No

P/IB 03/05773

A. CLASSIFICATION OF SUBSTANCE
IPC 7 G06T7/00 MATTER G06T15/20 H04N7/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04N H04M G06K G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99/57900 A (MYERS JOHN KARL) 11 November 1999 (1999-11-11) page 1, line 9 - line 12 page 11, line 1 - line 31 page 31, line 22 - line 27 page 39, line 19 - line 22 page 52, line 17 - page 52, line 2 page 61, line 34 - page 62, line 2 figure 14 ----- -/--	1-23



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

8 December 2004

Date of mailing of the international search report

15/12/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Ellerbrock, T

INTERNATIONAL SEARCH REPORT

International Application No

P/IB 03/05773

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LI-AN TANG ET AL: "Automatic construction of 3D human face models based on 2D images" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP) LAUSANNE, SEPT. 16 - 19, 1996, NEW YORK, IEEE, US, vol. VOL. 1, 16 September 1996 (1996-09-16), pages 467-470, XP010202432 ISBN: 0-7803-3259-8 P467, left column P469, right column figures 8,9	
A	US 2002/106114 A1 (YAN JIE ET AL) 8 August 2002 (2002-08-08) page 1, paragraph 8 page 1, paragraph 11 page 4, paragraph 40	1-23

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No

/IB 03/05773

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
W0 9957900	A	11-11-1999	AU	4307499 A		23-11-1999
			WO	9957900 A1		11-11-1999
<hr/>						
US 2002106114	A1	08-08-2002	NONE			
<hr/>						

METHOD AND APPARATUS FOR CORRECTING A HEAD POSE
IN A VIDEO PHONE IMAGE

5 The present invention relates to video phone systems, and more particularly, to a method and apparatus for correcting a head pose in a video phone image.

The consumer marketplace offers a wide variety of communications and media options. For example, various video phones are known that enable audio and video communications between users connected over a telephone line. A video phone system typically includes a microphone and speaker for enabling bidirectional audio communications and a camera
10 and display for enabling bidirectional video communications.

The technology for video phone applications has advanced to a point where video phone options are now being offered by many wireless telephone service providers. Wireless video phones thus enable audio and video communications between users connected over a wireless link. One common problem with video phone communications that is particularly
15 problematic with mobile users is that one or both participants to a video phone call may not be able to present a frontal face image to the camera at all time. For example, if a user is walking and looking at the sidewalk while holding the camera portion of the video phone in his or her hand, then the remote participant will typically see a "chin view" of the user's face. Similarly, if a user is sitting at a desk and turning his or her head to look at a
20 computer display, while the camera portion of the video phone is positioned on the user's desk, then the remote participant may see a "profile view" of the user's face.

A need therefore exists for a method and apparatus that correct a head pose in a video phone image, so that the remote participant will see a proper frontal view of the other participant. A further need exists for an improved technique for estimating and correcting
25 a head pose that is suitable for implementation on a wireless phone.

Generally, an image processing system and method are disclosed for correcting a head pose in a video phone image, so that a frontal view is presented on a display. A disclosed head pose corrector estimates the orientation of a head pose and adjusts the orientation of the head pose, if necessary, to present a frontal view. The orientation of the head pose is
30 adjusted by generating a three dimensional model of the face surface and adjusting the orientation of the three dimensional face model to provide the desired frontal view.

The disclosed head pose corrector may be included in a video phone of a user, to correct the head pose of transmitted or received images (or both) or may be included in a server on a network to automatically adjust the head shots of one or more participants to a video

phone communication. The computational requirements of the head pose corrector are suitable for implementation on a wireless video phone.

A more complete understanding of the present invention, as well as further features and advantages of the present invention, will be obtained by reference to the following detailed

5 description and drawings.

FIG. 1 illustrates a conventional video phone system;

FIG. 2 illustrates a network environment in which the present invention can operate; and

FIG. 3 is a flow chart describing an exemplary implementation of the image correction process of FIG. 2.

10 FIG. 1 illustrates a conventional video phone system 100. As shown in FIG. 1, the exemplary conventional video phone system 100 includes a microphone 110, a speaker 120, a camera 130 and a display 140 for enabling audio and video communications between two or more users. The conventional video phone system 100 may be embodied as any available video phone system, such as those commercially available from Sony
15 Ericsson Mobile Communications. It is noted that the microphone 110, speaker 120, camera 130 and display 140 may be integrated in a single unit, such as a desktop phone, or may be embodied as two or more modular units, as would be apparent to a person of ordinary skill in the art. For example, the camera 130 and display 140 may be embodied as modular attachments to a conventional telephone having the microphone 110 and speaker
20 120. In one particular implementation, the conventional video phone system 100 may be embodied as the T68i video phone system with a camera attachment, commercially available from Sony Ericsson Mobile Communications.

FIG. 2 illustrates a network environment 200 in which the present invention can operate.

As shown in FIG. 2, a first video phone system 210 incorporating features of the present
25 invention communicates over a network 220 with one or more additional video phone systems, such as the video phone system 270. The network 220 may be embodied as one or more wired or wireless networks, or a combination of the foregoing. The first video phone system 210 may be embodied as a conventional video phone system, such as the video phone system 100 shown in FIG. 1, as modified herein to provide the features and
30 functions of the present invention. The additional video phone systems 270 may be a conventional video phone system or a video phone system incorporating features of the present invention.

According to one aspect of the present invention, the video phone system 210 includes a head pose corrector 250 that employs a head pose estimation and correction process 300, discussed further below in conjunction with FIG. 3. The head pose corrector 250 may be integrated with a conventional video phone system 100 in a single unit, such as a desktop
5 phone, or may be embodied as a modular attachment to a conventional video phone system 100, as would be apparent to a person of ordinary skill in the art.

While the head pose corrector 250 is implemented in the exemplary embodiment in the video phone 210 of the first user, to process images of the local user that are being transmitted for display to the second user, the head pose corrector 250 could alternatively
10 process images of the remote user(s) that are received from one or more additional video phone systems 270 for presentation to the user of the first video phone 210. In a further variation, the head pose corrector 250 can be implemented in a server on the network 220 by a service provider to automatically adjust the head shots of all participants to a video phone communication in accordance with the teachings of the present invention.

15 FIG. 3 is a flow chart describing an exemplary implementation of the head pose estimation and correction process 300. Generally, the head pose estimation and correction process 300 ensures that a video phone image is a proper frontal view of a user. The computational requirements of the head pose estimation and correction process 300 are suitable for implementation on a wireless phone.

20 As shown in FIG. 3, the head pose estimation and correction process 300 initially obtains a sequence of images from the camera of the video phone system 210 during step 310. Thereafter, the head pose estimation and correction process 300 estimates the head pose during step 320 using pattern recognition techniques, such as the classification techniques described, for example, in Y. Li, S. Gong, and H. Liddell, "Support Vector Regression and
25 Classification Based Multi-View Face Detection and Recognition," IEEE Conf. on Automatic Face and Gesture Recognition 2000, incorporated by reference herein. Generally, the classification technique employed during step 320 will provide a characterization of the head pose, such as frontal view, chin view or profile view. In one variation, the classification techniques also provide the extent to which a chin view or
30 profile view deviates from a true frontal view. While many methods for estimating a head pose are computationally intensive, and susceptible to noise, the present invention recognizes that a practical solution is obtained in a video phone environment where in most cases a facial image is expected.

A test is then performed during step 330 to determine if the head pose is a frontal view. If it is determined during step 330 that the head pose is a frontal view, then the head pose correction techniques of the present invention are not needed and the unmodified image is transmitted during step 340.

- 5 If, however, it is determined during step 330 that the head pose is not a frontal view, then a three dimensional model is computed of the face surface from a sequence of facial images during step 350 using computer vision techniques, such as "structure from motion" techniques. For a detailed discussion of suitable techniques for computing a three dimensional model of the face surface from a sequence of facial images, see, for example,
- 10 M. Brand, "Morphable 3D Models from Video," Computer Vision & Pattern Recognition (CVPR) (2001) or M. Brand, "Flexible Flow for 3D Nonrigid Tracking and Shape Recovery," Computer Vision & Pattern Recognition CVPR (2001), each incorporated by reference herein. While many methods for estimating a general surface are computationally intensive, and susceptible to noise, the present invention recognizes that a
- 15 practical solution is obtained in a video phone environment where in most cases a facial surface is expected.

- A test is then performed during step 360 to determine if the head pose is a profile view. If it is determined during step 360 that the head pose is a profile view, then symmetric facial assumptions are employed during step 370 to estimate the remaining portion of the head
- 20 that is not present in the profile view. Program control then proceeds to step 380.

If it is determined during step 370 that the head pose is not a profile view then the view must be a chin view or a forehead view and program control proceeds directly to step 380. During step 380, the orientation of three dimensional face surface is adjusted to provide a frontal view.

- 25 Specifically, the origin of the three dimensional face surface is moved from where the input images are taken to a point in front of the nose point of the face surface. For example, chin view images are taken from a point below the desire origin and therefore origin correction is achieved by moving the three dimensional coordinates upwards. Similarly, forehead view images are corrected by moving the three dimensional
- 30 coordinates downward. Profile view images are corrected by rotating the three dimensional coordinates of the face surface by 90 degrees above the vertical axis of the surface. A frontal view can then be obtained by applying a standard perspective projection.

The modified image is then transmitted to the remote user during step 390. Thereafter, program control terminates.

It is to be understood that the embodiments and variations shown and described herein are merely illustrative of the principles of this invention and that various modifications may be
5 implemented by those skilled in the art without departing from the scope and spirit of the invention.

CLAIMS:

1. A method for processing an image containing at least a portion of a head of a human in a video phone system, comprising:
estimating an orientation of said head in said image using a pattern recognition technique;
computing a three dimensional model of a face surface of said human using a computer vision technique; and
adjusting an orientation of said three dimensional face surface model to provide a frontal view.
2. The method of claim 1, wherein said computing step further comprises the step of using a symmetric face assumption to obtain a complete three dimensional face surface model for a profile view.
3. The method of claim 1, wherein said computing step further comprises the step of employing a structure from motion technique to obtain said three dimensional face surface model.
4. The method of claim 1, wherein said estimating step applies a classification technique.
5. The method of claim 1, wherein said computing step generates a morphable three dimensional model.
6. The method of claim 1, further comprising the step of mapping said three dimensional face surface model having an adjusted orientation to a two dimensional space.
7. The method of claim 1, further comprising the step of transmitting said adjusted image to a remote user.
8. The method of claim 1, further comprising the step of presenting said adjusted image to a local user.

9. An image processor for use in a video phone system, comprising:
a memory for storing an image containing at least a portion of a head of a human; and
a head pose corrector that (i) estimates an orientation of said head in said image using a pattern recognition technique; (ii) computes a three dimensional model of a face surface of said human using a computer vision technique; and (iii) adjusts an orientation of said three dimensional face surface model to provide a frontal view.
10. The image processor of claim 9, wherein said head pose corrector is further configured to use a symmetric face assumption to obtain a complete three dimensional face surface model for a profile view.
11. The image processor of claim 9, wherein said head pose corrector is further configured to employ a structure from motion technique to obtain said three dimensional face surface model.
12. The image processor of claim 9, wherein said head pose corrector is further configured to apply a classification technique to obtain said head orientation.
13. The image processor of claim 9, wherein said three dimensional face surface model is a morphable three dimensional model.
14. The image processor of claim 9, wherein said head pose corrector is further configured to map said three dimensional face surface model having an adjusted orientation to a two dimensional modified image.
15. The image processor of claim 14, wherein said two dimensional modified image is transmitted to a remote user.
16. The image processor of claim 14, wherein said two dimensional modified image is presented to a local user.
17. A video phone system, comprising:
a memory for storing an image containing at least a portion of a head of a human; and

a head pose corrector that (i) estimates an orientation of said head in said image using a pattern recognition technique; (ii) computes a three dimensional model of a face surface of said human using a computer vision technique; and (iii) adjusts an orientation of said three dimensional face surface model to provide a frontal view.

18. The video phone system of claim 17, wherein said head pose corrector is further configured to use a symmetric face assumption to obtain a complete three dimensional face surface model for a profile view.

19. The video phone system of claim 17, wherein said head pose corrector is further configured to employ a structure from motion technique to obtain said three dimensional face surface model.

20. The video phone system of claim 17, wherein said head pose corrector is further configured to apply a classification technique to obtain said head orientation.

21. The video phone system of claim 17, wherein said head pose corrector is further configured to map said three dimensional face surface model having an adjusted orientation to a two dimensional modified image.

22. The video phone system of claim 21, wherein said two dimensional modified image is transmitted to a remote user.

23. The video phone system of claim 21, wherein said two dimensional modified image is presented to a local user.